

presence of a motif approximately -300 bp upstream of the transcription start point, called the endosperm box, was responsible for endosperm-specific expression. The endosperm box was subsequently considered to consist of two different motifs: the endosperm motif (EM) (canonical sequence TGTAAG) and the GCN 4 motif (canonical sequence (SEQ ID NO: 18) G/ATGAG/CTCAT). The GCN4 box is considered to regulate expression according to nitrogen availability (Muller and Knudsen, 1993). The *wSBE I-D4* promoter contains a number of imperfect EM-like motifs at approximately -100, -300 and -400 as well as further upstream. However, no GCN4 motifs could be found, which lends support to the idea that this motif regulates response to nitrogen, as starch biosynthesis is not as directly dependent on the nitrogen status of the plant as storage protein synthesis. Comparison of the promoters for *wSBE I-D4* and *D2* (Rahman *et al.*, 1997) indicates that although there are no extensive sequence homologies there is a region of about 100 bp immediately before the first encoded methionine where the homology is 61% between the two promoters. In particular there is an almost perfect match in the sequence over twenty base pairs (SEQ ID NO:19)CTCGTTGCTTCC/TACTCCACT, (positions 4723-4742 of the *wSBE I* sequence), but the significance of this is hard to gauge, as it does not occur in the rice promoter for SBE I. The availability of more promoters for starch biosynthetic enzymes may allow firmer conclusions to be drawn. There are putative CAAT and TATA motifs at positions 4870 and 4830 respectively of *wSBE I-D4* sequence. The putative start of translation of the mRNA is at position 4900 of *wSBE I-D4*.

Please replace the paragraphs beginning on page 45 at lines 15, 24, 29 and 33 with the following rewritten paragraphs, respectively:

(SEQ ID NO: 20)

5' GCGGCCGCTC CCTGGCCGAC TTGGCCGAAG CTTGCATGCC TGCAGGTCGA
CTCTAGAGGA TCCCCGGGTA CCGAGCTCGA ATTCAATCGAT GATATCAGAT
CCGGGCCCTC TAGATGCGGC CGCATGCATA AGCTT 3'

(SEQ ID NO: 21)

5'....CGCGCGCCCA CACCCTGCAG GTCGACTCTA GAGGATCCAT GGTGAGCAAG 3'.